



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Measuring the Broadband Access Divide

Citation for published version:

Fourman, MP 2016, Measuring the Broadband Access Divide. in *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development (ICTD '16)*. ACM, pp. 1-4, Eighth International Conference on Information and Communication Technologies and Development, Ann Arbor, Michigan, United States, 3/06/16. <https://doi.org/10.1145/2909609.2909616>

Digital Object Identifier (DOI):

[10.1145/2909609.2909616](https://doi.org/10.1145/2909609.2909616)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Proceedings of the Eighth International Conference on Information and Communication Technologies and Development (ICTD '16)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Measuring the Broadband Access Divide

Michael Paul Fourman
School of Informatics
10 Crichton Street
Edinburgh, Scotland
michael.fourman@ed.ac.uk

ABSTRACT

Decreases in a Gini index for broadband uptake have been interpreted as evidence of a narrowing digital divide. Nevertheless, a significant divide persists.

We propose two related indices, both well-known in the study of health inequalities (Wagstaff *et al.* 1991, 2005), as measures for the depth and breadth of the digital divide. Our concern is the contribution of the digital divide to social inequalities and cycles of deprivation. Depth quantifies the barriers to digital inclusion presented by existing deprivation. Breadth provides a measure of the degree to which the digital divide tends to reinforce existing inequalities.

Using data for broadband uptake in Scotland, we show how breadth and depth can be used, at local scale, to plan and assess interventions intended to close the divide. We also analyse global data from the International Telecommunications Union (ITU). The breadth of the global divide, which we interpret as its impact on inter-national inequalities, has increased steadily since 2000. The depth of the global divide, which we interpret as a relative measure of the barriers to entry facing those offline, fell from 2000 to 2011, but has risen annually since then.

CCS Concepts

•Social and professional topics → Broadband access;

Keywords

Gini, digital, divide, global, local, inequality, deprivation.

1. INTRODUCTION

In his millennial State of the Union address, President Clinton announced tax incentives intended, *to close the digital divide and open opportunity for our people.*

Opportunity for all requires something else today—having access to a computer and knowing how to use it. This means that we must close the digital divide...
Bill Clinton, 2000 [3]

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICTD '16 June 03-06, 2016, Ann Arbor, MI, USA

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4306-0/16/06.

DOI: <http://dx.doi.org/10.1145/2909609.2909616>

How should we measure the divide?

In 2002 a report from the U.S. Department of Commerce (DoC) adapted the standard methodology for assessing the distribution of income, to produce a *Gini Coefficient for Computer and Internet Use* [8]. This adaptation is identical to the **concentration index** of Wagstaff *et al.* 1991 [10].

Cumulative benefit (population online) is plotted against cumulative population, ordered by income, wealth or some other measure of deprivation, to give a *Lorenz Curve*. Fig. 1, using data for computer use in 1997 (*op. cit.* p.24 Fig 2-1), shows an example. Population is ordered by income.

Inequality is indicated by deviation from the equally-dashed diagonal line representing perfect equality, and measured as the difference between the areas above and below the Lorenz curve.¹ This difference is divided by the area, p , of the enclosing rectangle² to give a concentration coefficient of 19%.

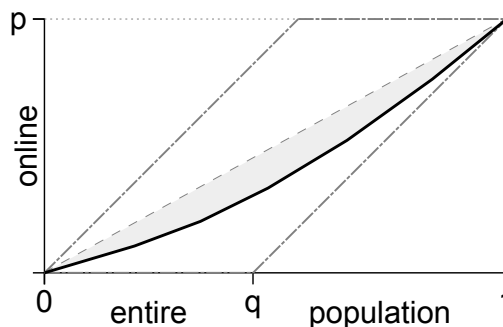


Figure 1: Concentration Index

The DoC report paints a rosy picture. Decreasing values of the index are interpreted to conclude, *inter alia*, that from 1984 to 2001 *the distribution of computers among households has moved continuously in the direction of less inequality.*

Others have used similar methods to reach similar conclusions. However, Sciadas [6] comments that, *The lowest income groups... continue to lose ground vis-à-vis the very high income groups*, whereas his computation of the Gini index, for the distribution of computer use against income, suggested that the digital divide was generally closing. Kelly [4], commenting on Cho's finding [2] of a reduction in the

¹This difference can be seen to be twice the shaded area between the Lorenz curve and the line of perfect equality.

²Here, p is the proportion of the population online; $q = 1 - p$ the proportion offline. The parallelogram is for later use.

global divide, says, *other evidence suggests that the progress in reducing the digital divide has occurred mainly as a result of middle-income countries catching up, whereas some of the least developed countries have actually been falling behind.*

The Gini index failed to capture perceived increases in the digital divide. Nevertheless, a 2009 United Nations report again used dramatic reductions in a Gini index, from 95% in 2000 to less than 50% in 2008, to justify a claim that, *Inequality is shrinking* ([7] p. 16).

Recent publications recognise a persistent divide. A 2015 report from UNESCO and the ITU [1] says, *the digital divide is proving stubbornly persistent in terms of access to broadband Internet.* A 2016 report from the World Bank [5] says, *digital divides persist across income, age, geography, and gender.* Neither mentions the Gini index.

Renormalisation.

In 2005, Wagstaff [9] observed that, when the advantage considered is binary—as broadband uptake is—the concentration index, C , a relative measure of inequality, must be renormalised. Wagstaff *et al.* 1991 [10] also introduced a generalised concentration index as an absolute measure of inequality. This must also be modified—in this case, scaled—for application to a binary advantage.

We will call the relative measure, D , the **depth** of the divide, and the absolute measure, B , its **breadth**—we explain these terms below. Each has a simple algebraic definition, in terms of C , the concentration index, and p , the proportion of the population enjoying the advantage. If $q = 1 - p$ is the proportion excluded, then,

$$D = C/q \quad B = 4Cp = 4Dpq \quad (1)$$

Fig. 2 shows data from the DoC report— p is home computer uptake (p.3 Fig. 1-1); C is the concentration index (p.87 Fig. 9-3)—together with our computed values for B and D . From 1990 to 2002, the concentration index, C , for

Y	1990	1994	1998	1999	2001	2002
p	15.9%	22.6%	36.6%	42.1%	51.0%	56.5%
C	40%	39%	31%	30%	26%	23%
B	25.4%	35.3%	45.4%	50.5%	53.0%	52.0%
D	47.6%	50.4%	48.9%	51.8%	53.1%	52.9%

Figure 2: Home computer uptake (from DoC data)

Households with a Computer plotted against Income, fell consistently. Our indices, B and D , tell a different story.

Outline.

We will justify and demonstrate use of breadth and depth as measurements for the digital divide. In § 2 we show that both breadth and depth arise as natural measures of the effects of the divide on interactions between social atoms. We discuss how these measures can be used to assess progress and inform policies intended to reduce the digital divide.

In § 3.1 we use postcode-level data for Scotland to relate digital exclusion to the Scottish Index of Multiple Deprivation (SIMD). In § 3.2 we apply these indices to ITU data, and discuss their interpretation in that context.

2. QUANTIFYING THE DIGITAL DIVIDE

Digital inclusion has the potential to provide increased opportunities in health, education, social inclusion, and well-

being to individuals in all sectors of society. However, many factors of deprivation constitute barriers to digital inclusion. So the benefits of increasing inclusion often serve to reinforce and reproduce existing inequalities, by widening the opportunity gap. To assess the social impacts of increasing digital inclusion, we will quantify these effects.

Abstractly, we consider the effects of some binary advantage on the relationships between individuals from a population subject to some deprivation ordering, \prec , where $b \prec a$, (b is below, and a above), if b is more deprived than a .

Concretely, our individuals are households who may be online or offline. For each offline-online pair, (u, v) , of households, if the offline household is inferior ($u \prec v$), then the digital gap between these two households strengthens v 's superiority. On the other hand, if $v \prec u$, then v 's digital advantage provides opportunities that serve to reduce the existing inferiority.

We divide the set of all offline-online pairs into S , those that strengthen deprivation, and R , those that reduce it.

$$\begin{aligned} S &= \{(u, v) \mid u \text{ is offline, } v \text{ is online, } u \prec v\} \\ R &= \{(u, v) \mid u \text{ is offline, } v \text{ is online, } v \prec u\} \end{aligned} \quad (2)$$

If the distribution of broadband uptake were independent of deprivation, we should expect these two sets to have the same size. In general, wherever the dependence of uptake on deprivation has been studied, S is larger than R .

The excess of S over R provides a natural measure of deprivation dependence. To give a normalised index that is independent of the size of the population, we divide $(S - R)$ by the number of possible pairs, then scale to give an index that occupies the range $[-1, 1]$. Our two indices are defined by entertaining two different sets of possibilities. If N is the total number of individual households, we define,

$$D = \frac{S - R}{S + R} = \frac{S - R}{pqN^2} \quad B = \frac{S - R}{N^2/4} \quad (3)$$

The depth index considers only the offline-online pairs. The breadth index considers all pairs.

We will now show that these are precisely Wagstaff's indices (1). Consider again Fig. 1. The pecked lines along the top and bottom of the parallelogram represent the Lorenz curves for two extremely unequal distributions.³ In one extreme, represented by the lower line, each offline household would be more deprived than every online household. This Lorenz curve follows the horizontal axis through the offline population (of size $q = 1 - p$), and then rises, with slope 1, through the online population. The concentration coefficient for this distribution is $C = q$. The Lorenz line for the other extreme, in which the most deprived sections of the population are those online, traces the top of the parallelogram to give a coefficient $C = -q$. Wagstaff suggested the renormalisation $D = C/q$ to give an index that uses the full $[-1, 1]$ range. For this example, $D = 42.1\%$.

This renormalisation amounts to dividing the difference in areas above and below the Lorenz curve by the area of the parallelogram of Fig. 1, instead of the area of the rectangle. For a point (x, y) on the Lorenz curve x is cumulative population, and y cumulative online population. We transform the parallelogram to a rectangle, and represent

³In September 1997, 54% of individuals age 3 and over were computer users. The values $p = 0.54$ and $q = 0.46$ represent the proportions of the population online and offline.

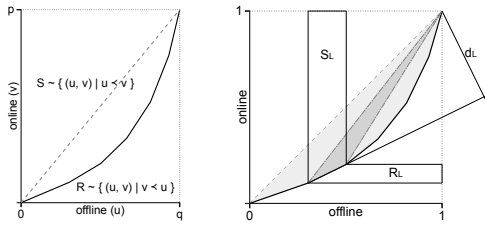


Figure 3: Depth of the Divide

the same curve on a plot of cumulative online population, $v = y$, against cumulative offline population, $u = x - y$.

This is shown on the left-hand side of Fig. 3. The rectangle now represents the set of offline-online pairs, (u, v) , of households, sorted in each dimension by our deprivation ordering, \prec . The definition of the Lorenz curve means that it separates the pairs in R , above, for which $u \prec v$, from those in S , below, for which $v \prec u$.

Let S_u, R_u be the numbers of online households above and below u in the deprivation ordering. We define the deprivation $\Delta(u)$ of an offline household, u , relative to the online population, by:

$$\Delta(u) = (S_u - R_u)/P \quad (4)$$

Since $S = \sum_u S_u$ and $R = \sum_u R_u$, where we take these sums over all offline households, the **depth** index represents the expected level of deprivation of an offline household. In so far as the various factors of deprivation act as barriers to digital inclusion, this provides a measure of the obstacles that must be overcome to get each offline individual online. The DoC data shows that those who remained offline were increasingly those who faced the highest barriers.

The **breadth** of the divide is an absolute measure of the degree to which the digital divide acts to strengthen existing divides. The generalised concentration index compares $S - R$ with the total number of pairs N^2 . Thus it measures the net effect of the digital divide on all possible binary interactions.

An extreme case for this index occurs when the more deprived half of the population is offline and the less deprived is online (or vice-versa). We have $P^2/4$ offline-online pairs, and they all fall in the set S where digital disadvantage acts to strengthen (or reduce) existing deprivation. Thus, if we apply Wagstaff's generalised concentration index to a binary advantage, the factor of 4 is required to give an index that exploits the full $[-1, 1]$ range of possible values.

For the DoC data presented earlier, the growing breadth of the divide is an indication of the increasing societal impact of the growing contribution of the digital divide to reinforce existing inequality.

The right-hand graph of Fig. 3 shows again the same curve scaled to the unit square. This has all the advantages of the traditional Gini plot. The depth index is represented as twice the area between the Lorenz curve and the line of perfect equality, and we can plot and compare curves for different populations and different years on the same diagram.

We can also use this diagram to quantify the distribution of inequality. The curve is made up of line segments, L , each representing a segment of the population. The net weight of digital disadvantage on one segment of the population is represented by a difference: the area above it, S_L , representing the pairs in S whose offline member is in L , minus

the area to its right, R_L , representing the pairs in R whose online member is in L . The area of the shaded triangle is half of this difference. The depth of L 's digital disadvantage is represented by d_L , the height of the triangle.

2.1 Related work

Breadth and depth are directly related to the Mann-Whitney-Wilcoxon rank-sum. The decile dispersion measure is closely related to our presentation of a Lorenz curve for Wagstaff's index. It tells us the ratio of the slopes of the starting and ending segments of the curve, but it ignores the middle ground.

The Theil index is a well-known information-theoretic measure of dispersion, and hence inequality. However, it cannot capture the links between advantage and deprivation that we want to measure.

3. EXAMPLES

We have already seen that the simple definitions (1) of B and D make it easy to revisit earlier studies and compute values for breadth and depth. Here we briefly describe two examples, to show how we can extract more information from primary data.

3.1 Scotland's Divide

Detailed data on broadband connections is recorded by service providers, for their own business purposes. In the UK, Ofcom has recently started to publish data giving the number of broadband connections in each postcode. Each of the five national ISPs, who together cover 90% of domestic connections, provides data which is then aggregated by Ofcom. We have combined this at postcode-level with data for the Scottish Index of Multiple Deprivation (SIMD), and census data (giving numbers of households). The data set covers around 180K postcodes, including a total of 2.5M households, over three successive years, 2013–2015.⁴ Fig. 4 shows that uptake has risen; the divide has become narrower, but deeper.

The maps show the relative depth of disadvantage (darker is deeper) experienced in each of Scotland's 32 local authority areas, in 2013 on the left, and 2015 on the right. We see that the position of the Western Isles, relative to rest of Scotland, has improved. However, those offline in East Renfrewshire, Glasgow City, Argyll and Bute, and the Western Isles suffer a disproportionate disadvantage.

Maps showing the distribution of the depth of the digital divide over smaller areas, provide a fine-grained analysis that suggests loci for targeted interventions.⁵

3.2 The Global Divide

To compute breadth and depth we require data on numbers of connections and numbers of households. We use ITU figures for broadband uptake. Numbers of households are computed, by division, from World Bank Total Population data⁶, combined with household size data for 68 countries from 2000–2012 assembled and interpolated by TekCarta⁷,

⁴This work contains public sector information licensed under the UK Open Government Licence v2.0.

⁵<http://idea.ed.ac.uk/digiscot>

⁶<http://data.worldbank.org/indicator/SP.POP.TOTL>

⁷<http://www.generatorresearch.com/tekcarta/databank/households-average-household-size/>

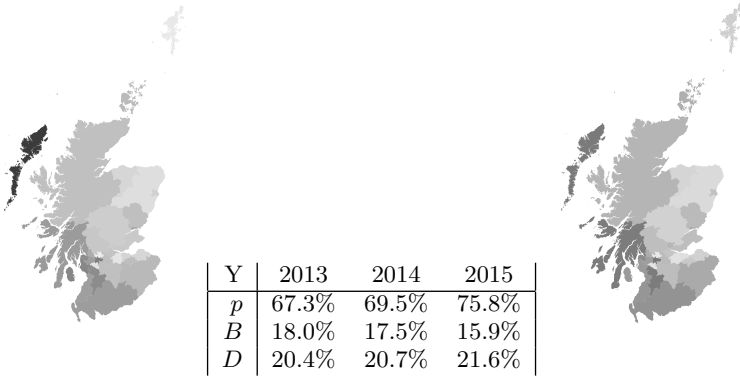


Figure 4: Households connected in Scotland

which we have extrapolated to 2013-14. We order these countries by level of broadband uptake (per household) and plot cumulative proportion of online households against cumulative proportion of offline households.

Fig. 5 shows depth, D , breadth, B , and uptake, p , for households connected, across the 68 countries for which we have been able to determine these figures for the years 2000–2014. We see that the depth of the divide reduced annually in the period 2000–2011, but has increased since then. Meanwhile, the breadth of the divide has steadily increased.

We believe these figures provide a lower bound for the global digital divide. They ignore many, mainly poorly connected, countries for which we have no data, and they ignore within-country inequalities.

Our Lorenz curve for each year is also drawn. Five years are highlighted, in bold in the table, and with thicker strokes for successive years. The curve for 2000 shows the greatest depth; the curve for 2011 shows the least.

Each successive curve from 2001 to 2004 dominates the 2000 curve and its predecessors. The curves for 2005–2007 show a decrease in D , and a changing pattern of inequality, with no Lorenz domination. From 2007–2011 we again see successive curves that dominate their predecessors. Finally, from 2011 to 2014 inequality increases, as more people go online in the more connected countries.

In the 2014 curve we see three clear groups of countries. Our Lorenz curve is approximated by three straight line segments. Their different slopes show different levels of opportunity. The first section, which includes India, for example, accounts for 40% of the offline households, but less than 5% of those online. The second section, dominated by China, is fast catching up with the digital leaders. The leaders, in the final section, include the USA and most of Europe.

Roughly 50% of the online households are in a group of well-connected countries that accounts for only 10% of the offline households. In these countries the odds of being online are over 3 : 1. The next 45% of the online households are found in a group of moderately-connected countries that accounts for around 40% of those offline. Their odds of being online are roughly 3 : 4. The final, poorly connected group includes around 5% of the online households, and 45% of those offline. In one of these countries, your odds of being online are roughly 3 : 40.

The 2015 State of Broadband report [1] says, *Network effects and externalities that multiply the impacts of ICTs require minimum adoption thresholds before those impacts can begin to materialize*, and suggests that, *multiplier effects may be widening the overall digital divide at a greater rate*

The digital divide across 68 countries, from ITU Fixed Broadband data, 2000 - 2014.

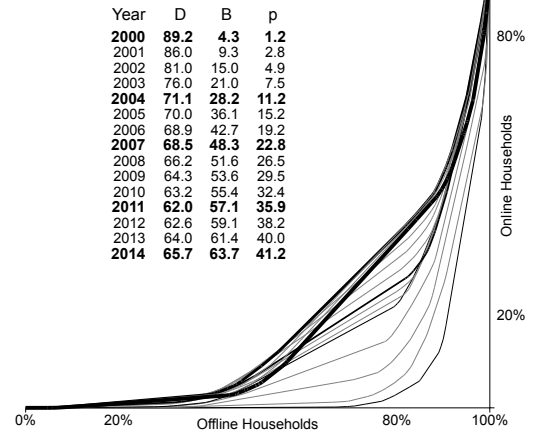


Figure 5: The Global Digital Divide 2000-2014

than simple adoption numbers suggest. Ordering countries by rate of uptake means that our depth measure captures such multiplier effects.

We interpret the recent deepening of the global divide as an indication that those nations still not online face increasing barriers to entry to the digital economy.

4. CONCLUSION

Breadth and depth provide better measures than the Gini, of our sometimes faltering progress on digital inclusion.

5. REFERENCES

- [1] BROADBAND COMMISSION FOR DIGITAL DEVELOPMENT. The state of broadband, 2015.
- [2] CHO, C. How to measure the digital divide? ITU/KADO Symposium on Building Digital Bridges, September 2004.
- [3] CLINTON, W. J. Address before a joint session of the congress on the state of the union, January 2000.
- [4] KELLY, T. Twenty years of measuring the missing link. In *Maitland+20: Fixing the missing link*, G. Milward-Oliver, Ed. Anima Centre Limited, Bradford on Avon, 2005, pp. 23–33.
- [5] MISHRA, D., DEICHMANN, U., ET AL. World development report: digital dividends, 2016.
- [6] SCIADAS, G. Unveiling the digital divide. Statistics Canada, October 2002.
- [7] UN ICT ANALYSIS SECTION. Information economy report 2009: Trends and outlook in turbulent times. UN Conference on Trade and Development, 2009.
- [8] VICTORY, N. J., AND COOPER, K. B. A nation online: How Americans are expanding their use of the Internet, February 2002.
- [9] WAGSTAFF, A. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics* (2005), 429–432.
- [10] WAGSTAFF, A., PACI, P., AND VAN DOORSLAER, E. On the measurement of inequalities in health. *Social Science & Medicine* 33, 5 (1991), 545 – 557.